BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 73(2), 2025, Article number: e153229
DOI: 10.24425/bpasts.2025.153229

ARTIFICIAL AND COMPUTATIONAL INTELLIGENCE

Machine learning-based fault detection in transmission lines: A comparative study with random search optimization

Yıldırım ÖZÜPAK*

Dicle University, Silvan Vocational School, Electrical Department, Diyarbakır, Türkiye

Abstract. Regular and fast monitoring of transmission line faults is of immense importance for the uninterrupted transmission of electrical energy. Rapid detection and classification of faults accelerate the repair process of the system, reducing downtime and increasing the efficiency and reliability of the power system. In this context, machine learning stands out as an effective solution for transmission line fault detection. In this study, fault detection is performed using machine learning techniques such as decision trees, logistic regression, and support vector machines. Random search hyperparameter optimization was applied to improve the performance of the models. The models were trained and tested with data from fault-free and faulted cases. While the support vector machines model showed the lowest performance with 74.19% test accuracy, the logistic regression model achieved 97.01% test accuracy. The decision tree model showed the best performance with low error rates. Error measures such as root mean square error (RMSE) and mean absolute error (MAE) were also used to evaluate the predictive power of the models. This research demonstrates how machine learning-based methods can be effectively used in the detection of transmission line faults and presents the performance of different algorithms in a comparative manner.

Keywords: power systems; fault detection; transmission line; machine learning; random search; regression.

1. INTRODUCTION

To meet the increasing demand for electricity, the electric power system is growing and becoming overall more complicated. The distribution, transmission, and generating systems constitute the three primary parts of the electric power system. In addition to distributing electricity to nearby customers, the system transfers electrical energy from producing facilities to regional substations, often via high voltage.

One of the biggest risks to the power system continuity and dependability is transmission line failure [1]. To avoid power system disruptions, it is critical to identify and eliminate these flaws. Faults in transmission lines can arise suddenly for several causes. Lighting, partial discharges (corona), wind, falling trees, ice and snow loading, polluted insulators, and pierced or damaged insulators are common sources of overhead line faults. Series and shunt faults are the two main categories into which transmission line defects fall. There are two types of shunt faults: symmetrical and asymmetrical. Single line-to-ground faults (A-G, B-G, C-G), line-to-line faults (AB, BC, CA), and double line-to-ground faults (AB-G, BC-G, CA-G) are all considered asymmetrical faults. Symmetrical faults include triple line-to-ground faults (ABC-G) and three-phase line faults (ABC) [2]. The categorization of transmission line fault types is shown in Fig. 1.

To stop electrical system damage and preserve electrical power flow, protection techniques must quickly identify and categorize defects. This accelerates the correction of undesired power losses, protecting linked equipment. To preserve power flow and restore the electrical system stability, it is critical to identify the precise site of the fault [3]. For an exceptionally long time, line identification, classification, and localization techniques have been effectively used by researchers. The assessment and utilization of several fault analysis techniques by electricity distribution firms facilitate the identification, classification, and forecasting of fault locations on transmission lines within power networks. As a result, they must address the challenging issue of deciding on a particular fault classification strategy. In this paper, a detailed review of various techniques used for precise fault identification, categorization, and troubleshooting on transmission lines in electric power systems is presented.

The study aims to emphasize the importance of using effective and innovative methods such as machine learning (ML) methods in the detection and classification of faults in transmission lines. In this context, the paper presents machine learning-based fault detection models specifically for electric power transmission line systems.

For the reliability of power systems, the rapid detection and classification of faults in transmission lines is critical to maintain the stability of the system. Where traditional methods are inadequate, solutions offered by artificial intelligence and machine learning techniques prove to be more effective, especially in the early detection and classification of symmetric and asym-

*e-mail: yildirim.ozupak@dicle.edu.tr

Manuscript submitted 2024-09-04, revised 2024-10-11, initially accepted for publication 2024-12-19, published in March 2025.

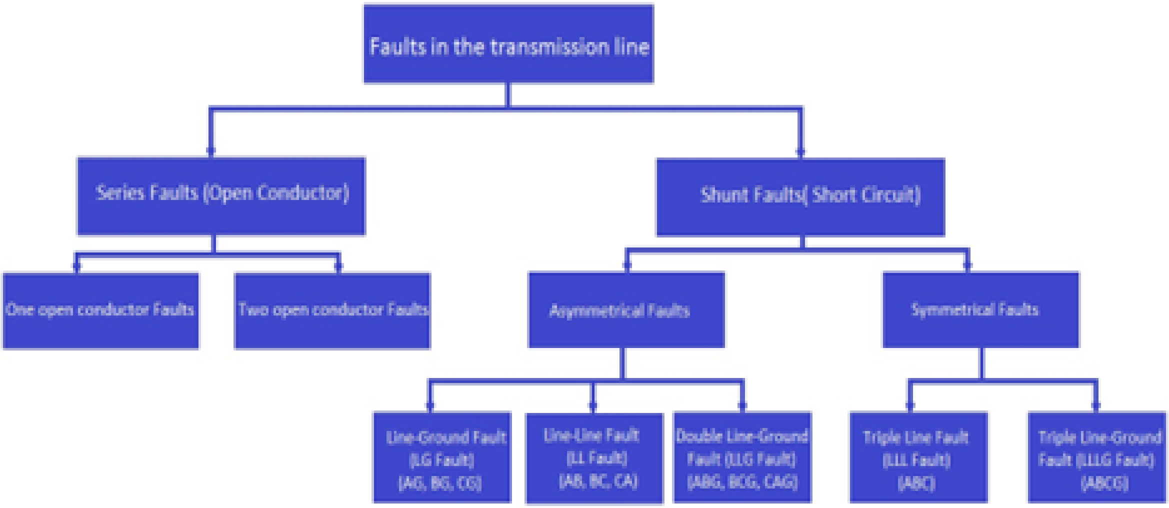

Fig. 1. Classification of fault types in transmission line [3]

metric faults. In this context, the study comparatively examines various machine learning models such as decision trees, logistic regression and support vector machines to detect and classify faults in transmission lines. The limitations of the study are that the dataset used is limited to certain types of faults and the applicability to a wider range of faults in real-time applications should be investigated in future work.

A brief overview of the issue is provided in Section 1 of this paper. The study is divided into the following sections. The research conducted for transmission line fault detection is included in Section 2. The materials and techniques required for the applications are presented in Section 3, and the results and comments are included in Section 4. Future work and conclusions are outlined in Section 5.

2. LITERATURE REVIEW

Considerable progress has been made in several areas in recent years regarding the identification and categorization of power system failures. Finding power system flaws is crucial, particularly in businesses where prolonged blackouts can result in significant financial losses. Historically, eye examination and trial-and-error switching were used by power companies as conventional means of defect identification. Furthermore, systems based on artificial neural networks were employed for defect categorization and detection.

Neural network techniques based on artificial intelligence (AI) are commonly employed to develop tools for investigating power system faults. The architecture on which ML is based comprises three layers: input, output, and hidden layer. The capacity of ML to learn on its own with only a few parameter adjustments required is one of its main advantages. The network modifies its weights during training, facilitating the implementation of real-time problems. It does have several drawbacks, though, such as the lengthy training periods required to analyze multidimensional situations. Large data sets and dispersed data are also necessary for ML to update the weights in its structural model, and multidimensional analysis will always need lengthy training times [4]. Failure analysis using artificial neural networks may be done in many ways. For instance, a model for defect classification and identification utilizing a feed-forward artificial neural network is shown in [5]. For each of the eleven fault situations, three-phase root mean square voltages and currents are used to train the model. This approach produces four fault parameters, which are used to locate and classify the fault, locate the ground fault, and indicate the fault presence.

The contribution of the study to the literature is provided by a comprehensive review of a number of artificial intelligence-based methods used in the detection and classification of transmission line faults. Through a comparative evaluation of the effectiveness of traditional and modern methods such as support vector machines, logistic regression and decision trees, it contributes to a better understanding of existing methods and techniques in this field. There are similar studies in the literature. However, the accuracy achieved with our proposed method surpasses that of other studies in the literature.

For defect identification and classification, Kumar *et al.* employed a feed-forward neural network trained with a back propagation technique [6]. Gowrishanka *et al.* classified and detected faults in transmission lines using artificial neural networks and discrete wavelet transforms [7]. The authors employ wavelet transform to extract information from the transient signal in both the time and frequency domain, and they apply ANN to classify faults. A neural network-based high-speed fault localization and detection approach was proposed by [8]. Albatsh *et al.* determined the weakest or most susceptible areas of the power system using four voltage stability indices. They conducted a comparison analysis of the indices based on how sensitive these indices were to the voltage collapse point [9]. The objective of Choudhury *et al.* 3 is to illustrate how a power grid experiencing a phase-ground failure differs from a system using solely PSS in terms of the coordinated functioning of PSS and SVC [10]. In their investigation, Shekhawat *et al.* highlighted significant concerns about voltage stability indices [11]. Using MATLAB-PSAT to simulate the Uganda Power System, Kyomugisha *et al.* assessed voltage stability and related improvement methods [12]. A review of several methods for accurate defect diagnosis, categorization, and troubleshooting is provided by Kanwal *et al.* [13].

Modified IEEE 39-bus and IEEE 9-bus systems were used in the implementation of the Jiriwibhakorn *et al.* study. The outcomes of the ANFIS installation demonstrate that ANFIS has a high degree of accuracy when predicting CCT values. Furthermore, a comparison between the ANFIS findings and the commonly utilized ANN results shows that the ANFIS produces better results faster [14]. Goni *et al.* use the extreme learning machine (ELM) method to propose a self-activating fault detection and classification system in their work. MATLAB Simulink is used to simulate fault data [15]. Upadhyay *et al.* constructed a flaw detector and classifier using a feed-forward ANN with a backpropagation technique. Performance is evaluated using mean squared error (MSE) and the confusion matrix of the classifier and detector [16]. Using an extended ANN, Obi *et al.* located the issue [17]. Using the Global Positioning System (GPS) and the Global System for Mobile Communications (GSM), problem analysis was conducted in this manner. For training, testing, and validation, the ANN mean square error value satisfies the desired MSE. Using MRA, the Taguchi technique, and the discrete wavelet transform, an artificial neural network model is created. Using the variations in wavelet entropies of three-phase voltages, three-phase currents, and neutral currents, the study aims to identify, classify, and forecast the location of faults. Using the Taguchi approach, an orthogonal dataset was created to train the ANN [18]. In a study by Affijulla and Tripathy, the three-phase current signal of the transmission line was deconstructed to the fifth level of detail. Using a DWT-ANN approach, these data were used to extract features and complete fault classification and detection. The classification accuracy of this hybrid model was determined to be 90.60% [19].

Applications for transmission line monitoring can make use of ML, a crucial machine learning technology. A neural network can be trained using offline data, which can be extremely helpful in fault analysis and detection in power systems. ML can be used to resolve the issue of distance relays falsely tripping owing to over- or under-reach faults. According to one study, ML can provide improved zone access capabilities together with transmission line failure detection. An ML method for suppressing fault resistance on distance relays based on pattern recognition is given in the research [20]. It was found that the network adjusted to changes in the power system network following extensive training with a variety of failure patterns. The authors of the research employed a feed-forward back-propagation neural network to identify faults in a three-phase transmission line. They fixed issues that, in the event of relay and circuit breaker failure, may cut off a sizable portion of the system network from the power source. The implementation of artificial intelligence techniques in power system fault analysis and detection was encouraged by the heuristic procedure utilized by power system operators in fault analysis. A neural network approach was used to locate the power system malfunctioning component. To represent busbars, transformers, and transmission lines, a multilayer perceptron neural network and general regression are utilized. Network topological changes might be managed by the intended module without requiring the network to be retrained.

The study presents a comprehensive review of various AI-based methods used in the detection and classification of transmission line faults. In this way, the effectiveness of traditional and modern methods is comparatively evaluated. A comparative analysis of machine learning techniques such as support vectors was conducted in the study. This analysis provides a better understanding of the methods in the literature by comparing the success of these techniques in detecting and classifying transmission line faults.

3. MATERIAL AND METHOD

3.1. Model of the transmission line in the power system

We looked at a power system that is depicted in Fig. 2 and consists of a three-phase, 50 Hz, 154 kV transmission line that is connected to a power source on both ends. Distance relay protection is the foundation of the ML relay that is utilized. The model depicts a Türkiye transmission line. Figure 2 was adapted from Akhikpemelo *et al.* [21].

MATLAB software is used to model the transmission system and simulate different forms of faults, producing a fault data set for network testing and training. Table 1 lists the power system parameters that were utilized.

Table 1
Electrical parameters of the power system

Parameters	Value
Frequency	50 Hz
Generator output voltage	14.6 kV
Transmission line voltage	154 kV
Desired active power	100 MW
Reactive power value	27 MVAr

Line-ground faults (A-G, B-G, and C-G), double-line-ground faults (AB-G, AC-G, and BC-G), line-line faults (AB, AC, and BC), and three-phase faults (ABC) are among the several types of faults, simulated by altering the fault type constraint of faults module. Various fault data were produced for L-G, LL-G, LL, and LLLL-G problems by utilizing the transmission line system MATLAB software model. A multilayer feed-forward network was built initially. Appropriately, further hidden layers were added to test the network viability. To set up the network for training, the biases and weights of each network object were initialized. The training was performed in batch training mode, which is considerably quicker and yields lower errors. The quantity of validation checks, the size of the gradient performance, and the network performance were all significant factors during the training process. Now that the network has been trained and validated, it can be utilized to determine how the relay system

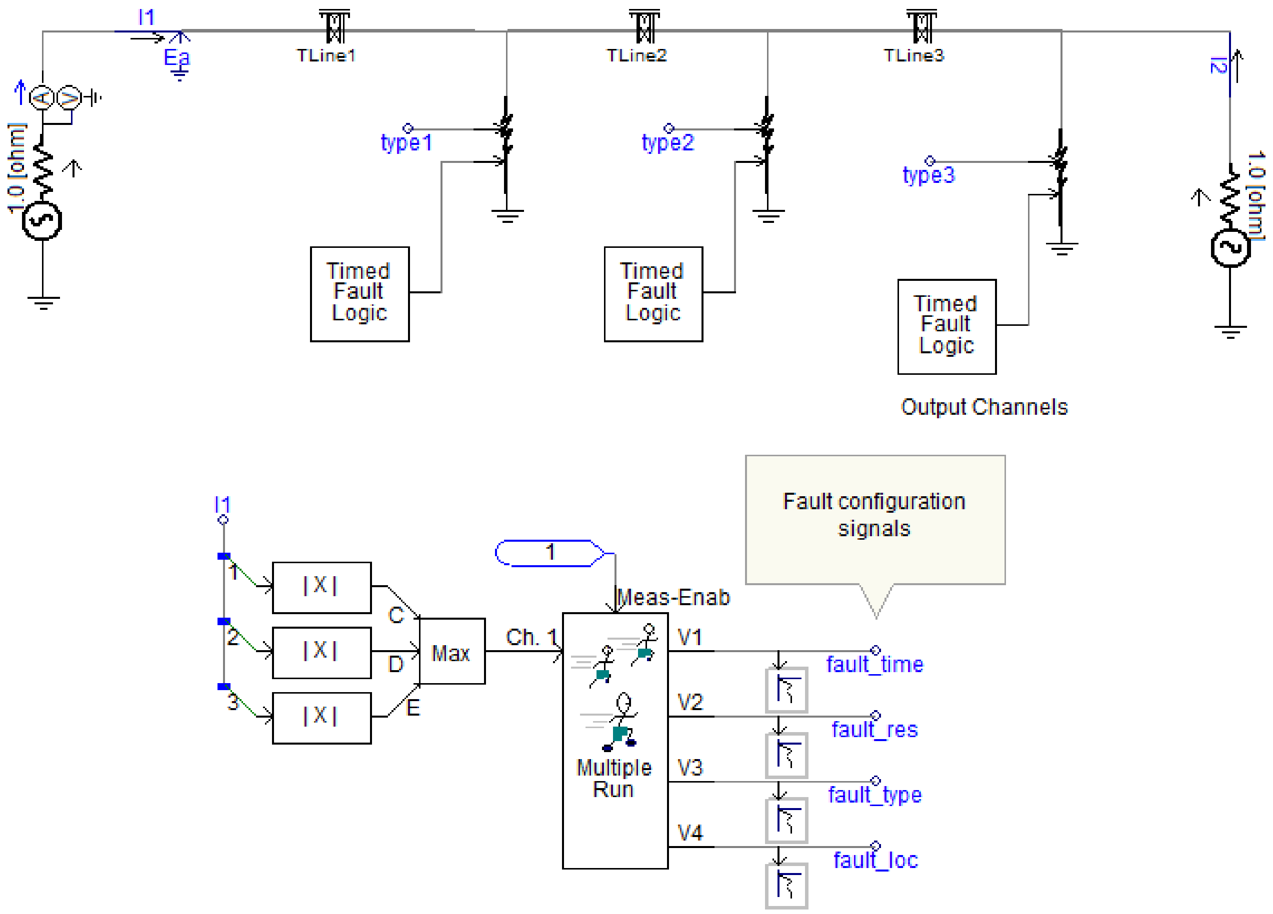

Fig. 2. Transmission line model [22]

will react to various failure scenarios. Then, rather than using the standard mathematical calculations, the neural network program was transferred to the microprocessor-based relay [23].

3.2. Support vector machines (SVM)

SVM is a machine learning technique used in high voltage transmission line fault detecting procedures. The SVM method is a useful tool for solving regression and classification issues. A hyperplane, or dividing line, is drawn in the space created by SVM, which represents data points in that space. The goal of this line is to increase the separation between the two groups. The points that are closest to this line are the support vectors, which establish the distance between classes [24].

3.3. Logistic regression

Logistic regression is a statistical modelling method used especially in classification problems. This regression is used for situations where the dependent variable is categorical. It is usually used for binary classification but can also be extended to multi-class classification problems [25].

3.4. Decision tree

Decision tree-based models for detecting faults in high-voltage transmission lines are often built using machine-learning techniques. These models are designed to perform tasks such as detecting, classifying, and potentially identifying the causes of faults.

3.5. Random search optimization

Random search optimization is an effective approach to hyperparameter optimization of machine learning models. It attempts to find optimal hyperparameter combinations by evaluating randomly selected points in the parameter space. Unlike systematic search methods such as grid search, random search tries a certain number of random combinations instead of exhaustively exploring all possibilities. This approach significantly reduces the computational cost, especially for models with large and complex hyperparameter spaces, and allows for satisfactory results in less time. Furthermore, random sampling increases the probability of exploring a larger region of the hyperparameter space, which increases the chances of reaching the global optimum [26].

Random search can be more effective when working with high-dimensional data sets and complex models. While grid search is more expensive in terms of time and computational resources because it must try all combinations, random search can achieve similar or better performance by trying only a certain number of combinations. Especially in practical applications such as machine learning-based power system fault detection, the flexibility and computational efficiency of random search offer significant advantages in model optimization. This method can increase the generalization ability of the model by reducing overlearning and improving performance metrics, thus playing a critical role in improving the reliability and accuracy of machine learning models.

3.6. Training and testing dataset

An artificial neural network model designed to use three different algorithms was selected [27, 28]. This dataset contains scaled current and voltage values of three phases. For the six inputs in the dataset, there are 11 different fault conditions and no-fault conditions. The truth table for these fault types is shown in Table 2.

Instantaneous voltages and currents for all three phases are scaled by the neural network in use for five distinct fault circumstances as well as the fault-free state. The data collection contains 7861 labelled data points. The fault type of these data indicates that there are 1134 A-B-G faults, 1133 A-B-C-G faults, 1129 A-G faults, one A-B-C fault, 1004 B-C faults, and 2365 No faults in the system.

The data set used in the study was obtained from simulations performed with MATLAB Simulink. Phase voltage and current values were collected for transmission line faults (e.g., A-G, A-B-G, B-C) and fault-free conditions. For each fault condition, the instantaneous current and voltage data of the phases were recorded and then normalized and added to the data set. The data set includes six input variables (voltage and current values of three phases) and 11 fault conditions and fault-free conditions. For example, for the A-G fault, the short circuit of phase A with the earth was simulated and the data in this case was labelled and recorded. 80% of the data set was used for training and 20% for testing. Machine learning models were trained and tested with these data.

Table 2
Truth table for various fault types

	Failure type	A line	B line	C line	Ground
1	A-G	1	0	0	1
2	B-G	0	1	0	1
3	C-G	0	0	1	1
4	A-B	1	1	0	0
5	A-C	1	0	1	0
6	B-C	0	1	1	0
7	A-B-G	1	1	0	1
8	A-C-G	1	0	1	1
9	B-C-G	0	1	1	1
10	A-B-C	1	1	1	0
11	A-B-C-G	1	1	1	1

To better understand the structure of the data in the dataset, the correlation between each feature was calculated and visualized as shown in Fig. 3. This correlation matrix shows the relation-

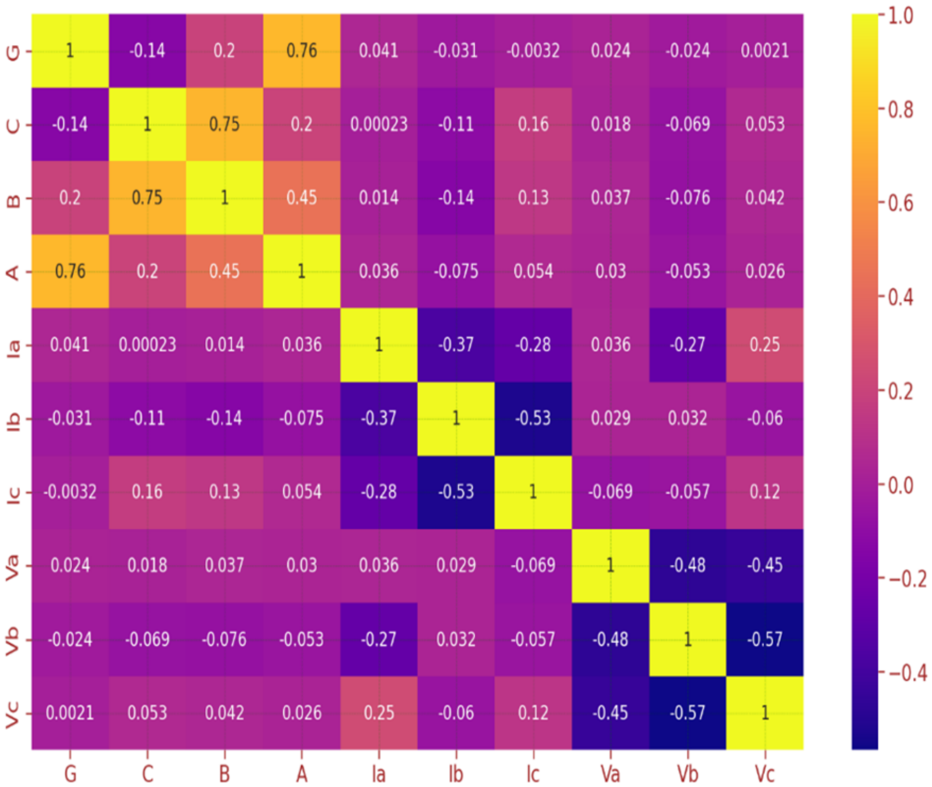

Fig. 3. Correlation matrix

ships between the various variables used for fault detection in the power systems considered in the study. The matrix contains the correlation coefficients between pairs of variables, ranging from −1 to +1. When the correlation coefficient is +1, there is a positive and strong relationship between the two variables, and when it is −1, there is a negative and strong relationship. A correlation coefficient of 0 indicates that there is no linear relationship between the variables. For example, a coefficient of 0.76 was calculated between G and A in the correlation matrix. This indicates that these two variables have a strong positive relationship. There is a coefficient of −0.37 between Ia and Ib. This indicates that there is a weak negative relationship between the two variables.

Such correlation relationships can be useful for understanding whether certain types of faults in the system are related to each other and for modelling faults more effectively. For example, highly correlated variables like A and G can be analyzed together, while negatively correlated variables like Ib and Ic should be considered for different analyses.

The correlation matrix plays a key role in fault detection because faults often represent deviations from normal in the system. These deviations can cause a significant change in correlation between different sensors. By identifying which sensors are more strongly correlated with each other, correlation matrix analysis can help pinpoint the location and type of a potential failure. A high correlation indicates that a deviation in a particular sensor may be associated with a similar deviation in other sensors, which could be a possible symptom of failure. Therefore, the correlation matrix can be used as an effective tool in fault detection.

It is anticipated that the mean square error number will drop as the neural network is trained. The validation check then looks at the deviations of the trained neural network. Both faulty and fault-free circumstances are identified by the fault detection approach. Using an 80% training and 20% test data split, a complete dataset comprising all simulated fault current and fault-free current data is produced.

4. RESULTS AND DISCUSSION

Python programming language was used for machine learning training and models were created with the widely preferred scikit-learn library. Scikit-learn provides a large set of tools that enable easy implementation of machine learning algorithms. Models such as decision trees, logistic regression, and support vector machines were trained and tested using this library. In addition, the random search method for hyperparameter optimization of the models was also implemented through the scikit-learn library. NumPy and Pandas libraries were used for data processing and analysis.

There are five different 'fault' and 'no fault' conditions in the neural network used. These states are A-G, A-B-G, B-C, A-B-C, A-B-C, A-B-C-G and No Fault. The current and voltage graphs with and without faults obtained with three different regression models are shown in the figures above. Logistic regression, support vector machines, and decision tree ML models were used. To improve the performance of the machine learning models used in this study, a hyperparameter optimization method such as random search is used. Random search optimization plays a significant role in maximizing the potential of the model. As a result of the application of random search optimization, the training and testing accuracy of the support vector machines, logistic regression, and decision tree models are improved, as well as the error metrics such as RMS and MAE. Especially in the decision tree model, this optimization algorithm was used to reduce the risk of overlearning and to increase the generalization ability of the model. Hyperparameter optimization with random search improves the performance of existing models and increases the reliability and success of these models in the application domain. This type of optimization contributes to the advancement of machine learning-based fault detection studies. The parameters used in the training with logistic regression are given in Table 3.

Table 3
Logistic regression parameters

Parameter	Value
Penalty	L2
Class_weight	none
Intercept_scaling	1
Max_iter	250
Random_state	50

As a result of the training in logistic regression, 97.1% training accuracy and 96.7% test accuracy were obtained in predicting the fault in transmission lines. RMSE is 0.691, MAE is 0.119, and R^2 Score is calculated as 85.181. The confusion matrix of the logistic regression is presented in Fig. 4. As can be seen in the confusion matrix, some A-G Fault is incorrectly predicted. This reduces the accuracy of the model.

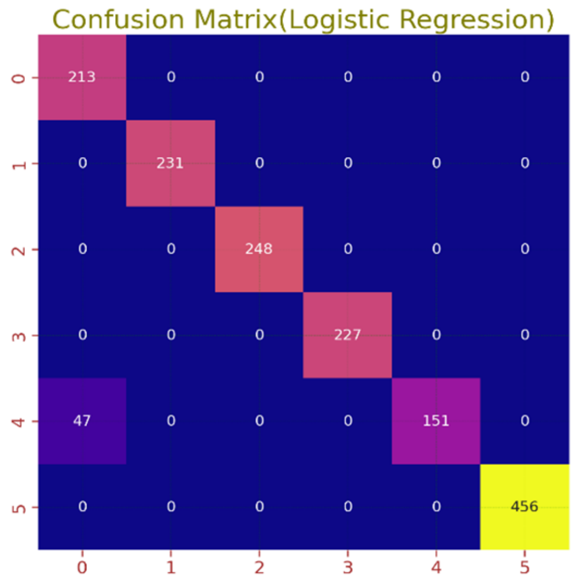

Fig. 4. Confusion matrix of logistic regression

The graph showing the prediction of transmission line fault and actual fault values predicted by the logistic regression algorithm is given in Fig. 5. The actual fault types in the data set and the fault type predicted by the logistic regression algorithm are given in Table 4.

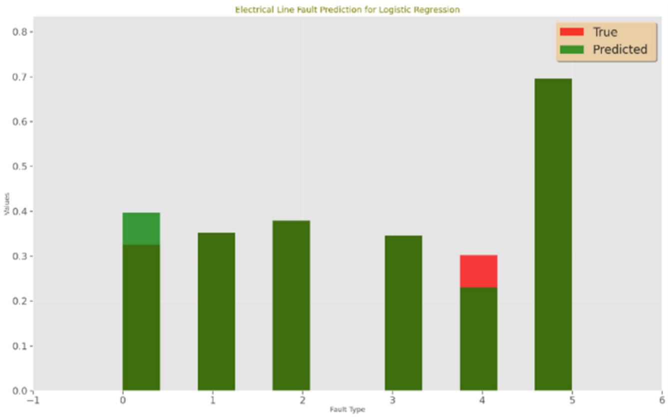

Fig. 5. Comparison of graphs of actual faults on the transmission line and faults predicted by logistic regression

Table 4
True and predicted values of ID numbers randomly extracted from the dataset

Dataset ID	True fault type	Predicted fault type
1556	2	2
3739	0	0
7451	5	5
4652	1	1
2594	4	0
4270	0	0
6816	5	5
1643	2	2
55	3	3
662	3	3

The parameters used in the training with support vector machines (SVM) regression are shown in Table 5.

SVM training resulted in 74.19% training accuracy and 76.51% test accuracy in predicting the fault in transmission lines. RMSE is 0.877, MAE is 0.388, and R^2 score is 76.136. The SVM confusion matrix is presented in Fig. 6. When the confusion matrix resulting from the SVM model is analyzed, it is seen that the predicted failure types are more inaccurate than logistic regression.

The graph showing the prediction of transmission line fault and actual fault values predicted by the SVM algorithm is given in Fig. 7. The actual faults in the data set and the faults predicted by the SVM regression algorithm are given in Table 6.

Table 5
SVM regression parameters

Parameter	Value
Epsilon	0.12
cache_size	200
Shrinking	true
Max_iter	–1
Coef0	0.0

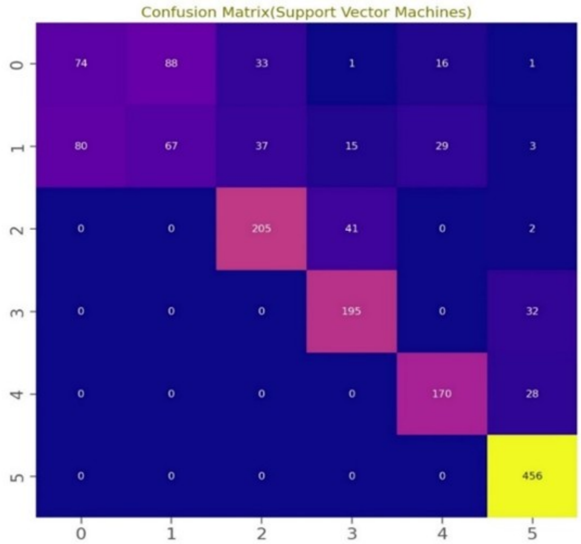

Fig. 6. Confusion matrix of SVM

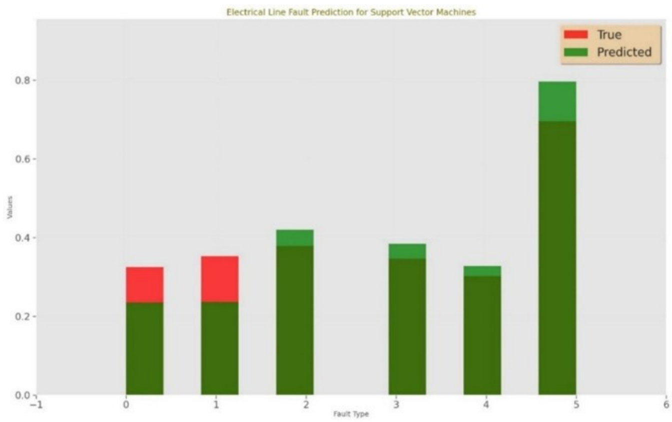

Fig. 7. Comparison of graphs of actual faults on the transmission line and faults predicted by SVM

The parameters used in the training with decision tree regression are shown in Table 7.

As a result of the training of the decision tree regression, 99.71% training accuracy and 99.87% test accuracy were obtained in predicting the fault in transmission lines. RMS is 0.017, MAE is 0.005, R^2 Score is 99.851. The confusion matrix of this regression is presented in Fig. 8.

Table 6
True and predicted values of ID numbers randomly extracted from the dataset

Dataset ID	True fault type	Predicted fault type
6816	5	5
6859	5	5
6731	5	5
3681	0	2
6326	5	5
1664	2	2
7053	5	5
5488	1	5
278	3	3
172	3	3

Table 7
Decision tree regression parameters

Parameter	Value
Splitter	best
Max_depth	none
min_samples_leaf	1
Random_state	none
min_weight_fraction_leaf	0.0

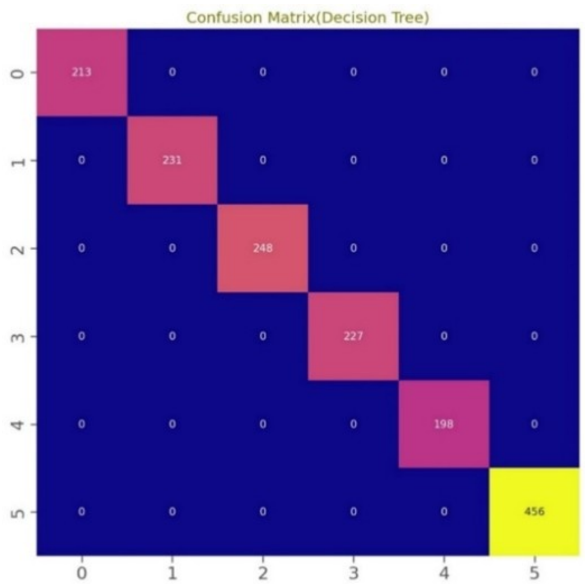

Fig. 8. Confusion matrix of decision tree

The graph showing the prediction of transmission line fault and actual fault values predicted by the decision tree algorithm is given in Fig. 9. The actual fault types in the data set and the fault types predicted by the decision tree algorithm are given in Table 8.

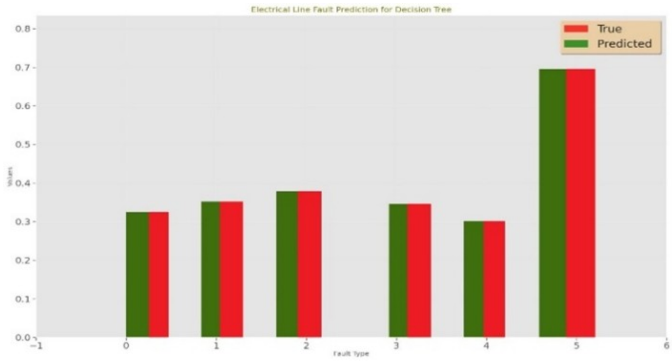

Fig. 9. Comparison of graphs of actual faults on the transmission line and faults predicted by decision tree

Table 8
True and predicted values of ID numbers randomly extracted from the dataset

Dataset ID	True fault type	Predicted fault type
3704	0	0
6525	5	5
2510	4	4
7172	5	5
311	3	3
4361	0	0
2221	2	2
129	3	3
4830	1	1
3459	0	0

Decision trees are used to classify data or perform regression analysis using a set of decision rules. This tree structure divides the dataset into small pieces and applies a specific decision algorithm to each piece. When building the decision tree, the algorithm tries to select features that best divide and classify the dataset. Each decision node is based on a specific feature and a specific value of that feature. In this way, the dataset is divided into smaller subsets.

Table 9 shows the performance of different regression algorithms. First, the support vector machines (SVM) model shows that the training accuracy is 76.51% and the test accuracy is 74.19%. This shows that the model provides sufficient learning on the training data but has limited generalization ability on the test data. High root mean square error (RMS) and mean absolute error (MAE) values can negatively affect the accuracy of the model predictions. The logistic regression model stands out with a training accuracy of 96.77% and a testing accuracy of 97.01%. This indicates that the model fits the dataset well and has high generalizability. The low RMS (0.691) and MAE

Table 9
Comparison of regression algorithms results

Regressor algorithm	RMS score	MAE score	R^2 score
Support vector machines	0.877	0.388	76.136
Logistic regression	0.691	0.119	85.181
Decision tree	0.017	0.005	99.851

(0.119) values support the success of the model, as the predictions are highly dependable. Finally, the decision trees model has the highest accuracy rates on training and test data (99.71% training, 99.87% test). It also shows a high R^2 score (99.851%) with extremely low RMS (0.017) and MAE (0.005) values. This suggests that both the learning and generalization ability of the decision tree model are excellent.

Overall, Table 9 reveals that decision trees perform the best, followed by logistic regression and SVM performs the worst. These results suggest that decision trees are the most suitable machine-learning model for transmission line fault detection. While logistic regression also provides a viable alternative, it is emphasized that the performance of SVM needs to be improved.

The study examines in detail various artificial intelligence-based methods used in the detection and classification of transmission line faults. Thus, the effectiveness of both traditional and modern methods can be comparatively evaluated. A comparative analysis of machine learning techniques such as support vector machines, logistic regression, and decision trees was conducted. This analysis provides a better understanding of the methods in the literature by comparing the success of these techniques in the detection and classification of transmission line faults.

The algorithm parameters used are optimized to improve the results. The proposed methods have higher accuracy rates than similar studies in the existing literature. In particular, the decision tree algorithm achieved 99.87% test accuracy in detecting and classifying errors as a result of optimization.

5. CONCLUSIONS

Disturbances in electric power transmission systems significantly impact reliability, often due to delays in fault identification and classification. These delays hinder prompt troubleshooting, emphasizing the need for early detection, rapid response, and swift restoration. This study explores the application of artificial neural networks for identifying and classifying faults in three-phase transmission lines. To address both symmetrical and nonsymmetrical issues, deep learning-based regression models, such as decision trees, logistic regression, and support vector machines, are employed. Random search optimization enhances the performance of these machine learning models. The models utilize phase currents and voltages from a Simulink/MATLAB transmission line model, with normalized instantaneous voltage and current inputs for fault detection and classification. Six distinct fault types are analyzed, demonstrating accurate identification of faults on 154 kV transmission lines. Validation with real-time data from Türkiye Transmission Company confirms the effectiveness of the model, with decision tree regression achieving 99.87% accuracy. All tested shunt faults were correctly identified and classified, highlighting the potential of ML-based models for improving system cost-effectiveness and maintenance efficiency. Future work will expand real-time applications, explore other fault types, and incorporate ensemble techniques combining machine learning and deep learning.

REFERENCES

[1] S.P. Ojha, B.K. Panigrahi, B. Sahu, and G. Dani, "Identification of double line to ground fault in electrical power transmission line using artificial neural network," *2nd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology*, 2022, pp. 1–5, doi: 10.1109/ODICON54453.2022.10010224.

[2] S. Kumar, A.K. Singh, and M.A. Kalam, "Fault prediction in transmission line using GOOSE protocol and artificial neural network," *Fourth International Conference on Electrical, Computer and Communication Technologies*, 2021, pp. 1–7, doi: 10.1109/ICECCT52121.2021.9616801.

[3] E.S. Prakash, "Electrical fault detection and classification," Kaggle, 2021. [Online]. Available: https://www.kaggle.com/datasets/esathyaprakash/electrical-fault-detection-and-classification (Accessed: Jan. 10, 2024).

[4] M.M. Taheri, H. Seyedi, and B. Mohammadi-ivatloo, "DT-based relaying scheme for fault classification in transmission lines using MODP," *IET Gener. Transm. Distrib.*, vol. 11, no. 11, pp. 2796–2804, 2017, doi: 10.1049/iet-gtd.2016.1821.

[5] A. Prasad, J.B. Edward, and K. Ravi, "A review on fault classification methodologies in power transmission systems," *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 1, pp. 48–66, 2018, doi: 10.1016/j.jesit.2017.01.004.

[6] S. Kumar, M. Swamy, and V. Venkatesh, "Artificial neural network based method for location and classification of faults on a transmission lines," *Int. J. Sci. Res. Publ.*, vol. 4, no. 6, pp. 1–6, 2014.

[7] M. Gowrishankar, P. Nagaveni, and P. Balakrishnan, "Transmission line fault detection and classification using discrete wavelet transform and artificial neural network," *Middle-East J. Sci. Res.*, vol. 24, no. 4, pp. 1112–1121, 2016, doi: 10.5829/idosi.mejsr.2016.24.04.23063.

[8] D.F. Leite, M.B. Hell, P. Costa, and F. Gomide, "Real-time fault diagnosis of nonlinear systems," *Nonlinear Anal.-Theory Methods Appl.*, vol. 71, no. 12, pp. 2665–2673, 2009, doi: 10.1016/j.na.2009.06.037.

[9] F.M. Albatsh *et al.*, "A comparative analysis of line stability indices for dynamic voltage stability," *International Conference*

on Engineering Technology and Technopreneurship, 2017, pp. 1–6, doi: 10.1109/ICE2T.2017.8215949.

[10] S. Choudhury, N. Khandelwal, and T.P. Dash, "A unified scheme of PSS and SVC for voltage profile improvement in electrical grid network," *1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology*, 2021, pp. 1–6, doi: 10.1109/ODICON50556.2021.9428989.

[11] N. Shekhawat, A.K. Gupta, and S. Kumar, "Voltage stability assessment using line stability indices," *3rd International Conference and Workshops on Recent Advances and Innovations in Engineering*, 2018, pp. 1–4, doi: 10.1109/ICRAIE.2018.8710435.

[12] R. Kyomugisha, C. Muriithi, and M. Edimu, "Voltage stability enhancement of the Uganda power system network," *2021 IEEE PES/IAS Power Africa*, 2021, pp. 1–5, doi: 10.1109/PowerAfrica52236.2021.9543385.

[13] S. Kanwal and S. Jiriwibhakorn, "Artificial intelligence-based faults identification, classification, and localization techniques in transmission lines – a review," *IEEE Latin Am. Trans.*, vol. 21, no. 12, pp. 1291–1305, 2023, doi: 10.1109/TLA.2023.10305233.

[14] S. Jiriwibhakorn, "Critical clearing time prediction for power transmission using an adaptive neuro-fuzzy inference system," *IEEE Access*, vol. 11, pp. 142100–142110, 2023, doi: 10.1109/ACCESS.2023.3341968.

[15] O.F. Goni *et al*., "Fast and accurate fault detection and classification in transmission lines using extreme learning machine," *e-Prime-Adv. Electr. Eng. Electron. Energy*, vol. 3, pp. 1–12, 2023, doi: 10.1016/j.prime.2023.100107.

[16] S. Upadhyay, S. Kapoor, and R. Choudhary, "Fault classification and detection in transmission lines using ANN," *2018 International Conference on Inventive Research in Computing Applications*, 2018, pp. 1029–1034, doi: 10.1109/ICIRCA.2018.8597294.

[17] E. Obi, A. Ezechukwu, and C.N. Ezema, "An extended ANN-based high-speed accurate transmission line fault location for double phase to earth fault on non-direct-ground," *Int. J. Eng. Sci. Technol.*, vol. 1, no. 1, pp. 31–47, 2019, doi: 10.29121/IJOEST.v1.i1.2017.04.

[18] Y. Hong and M. Cabatac, "Fault detection, classification, and location by static switch in microgrids using wavelet transform and Taguchi-based artificial neural network," *IEEE Syst. J.*, vol. 4, no. 2, pp. 2725–2735, 2019, doi: 10.1109/JSYST.2019.2925594.

[19] S. Affijulla and P. Tripathy, "A robust fault detection and discrimination technique for transmission lines," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6348–6358, 2017, doi: 10.1109/TSG.2017.2709546.

[20] A.P. Vaidya and A.V. Prasad, "Distance protection scheme for long transmission line protection considering the effect of fault resistance using ANN approach," *Int. J. Electr. Electron. Eng.*, vol. 1, no. 3, pp. 62–66, 2012, doi: 10.47893/IJEEE.2013.1044.

[21] Z. Bo Tang, "Adaptive partitioned random search to global optimization," *IEEE Trans. Autom. Control*, vol. 39, no. 11, pp. 2235–2244, Nov. 1994, doi: 10.1109/9.333768.

[22] A. Akhikpemelo, M. Evbogbai, and M. Okundamiya, "Fault detection on a 132 kV transmission line using artificial neural network," *Int. Rev. Electr. Eng.*, vol. 14, no. 3, pp. 220–225, 2019, doi: 10.15866/iree.v14i3.17054.

[23] E. Aslan, "Temperature prediction and performance comparison of permanent magnet synchronous motors using different machine learning techniques for early failure detection," *Eksploat. Niezawodn*., 2024, doi: 10.17531/ein/192164.

[24] A. Adamkiewicz and P. Nikończuk, "An attempt at applying machine learning in diagnosing marine ship engine turbochargers," *Eksploat. Niezawodn.*, vol. 24, no. 4, pp. 795–804, 2022, doi: 10.17531/ein.2022.4.19.

[25] M. Parsi, P. Crossley, P.L. Dragotti, and D. Cole, "Wavelet based fault location on power transmission lines using real-world traveling wave data," *Electr. Power Syst. Res.*, vol. 186, p. 106261, 2020, doi: 10.1016/j.epsr.2020.106261.

[26] E. Aslan, "Prediction and comparative analysis of emissions from gas turbines using random search optimization and different machine learning-based algorithms," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 72, no. 2, p. e151956, 2024, doi: 10.24425/bpasts.2024.151956.

[27] H. Sivaprasad, M.R. Lekkala, M. Latheef, J. Seo, K. Yoo, and C. Jin, "Fatigue damage prediction of top tensioned riser subjected to vortex-induced vibrations using artificial neural networks," *Ocean Eng*., vol. 15, no. 268, p. 113393, 2023, doi: 10.1016/j.oceaneng.2022.113393.

[28] E. Aslan and Y. Özüpak, "Detection of road extraction from satellite images with deep learning method," *Cluster Comput*. vol. 28, pp. 72–83, 2025, doi: 10.1007/s10586-024-04880.